Glycoconj J (2014) 31:259–269
DOI 10.1007/s10719-014-9519-9

Identification of glycosylated regions in pneumococcal PspA conjugated to serotype 6B capsular polysaccharide

Giovana Cappio Barazzone · Vittoria Pinto · Danilo Donnarumma · Martha Massako Tanizaki · Nathalie Norais · Francesco Berti

Received: 7 November 2013 / Revised: 26 February 2014 / Accepted: 3 March 2014 / Published online: 22 March 2014

Abstract Conjugate vaccines are being widely used since their introduction. Nowadays the interest in these vaccines is still growing and new antigens and conjugate chemistry are being studied and developed. Pneumococcal surface protein A (PspA) is one of the most studied pneumococcal antigens and is an important vaccine candidate. One approach to broaden the conjugate vaccine coverage could be the conjugation of the polysaccharide to a pneumococcal protein such as PspA. Previous results have shown that conjugated recombinant fragment of PspA (rPspA) not only maintained but also in some conjugates improved the induction of protective antibodies raised against the protein carrier. We describe here a characterization study to identify the domains of *Streptococcus pneumoniae* recombinant PspA (rPspA), from family 1 clade 1 and family 2 clade 3, involved in the conjugation with serotype 6B capsular polysaccharide.

Keywords Conjugate vaccines · *Streptococcus pneumoniae* serotype 6B · PspA · Mass spectrometry

Abbreviations

PspA	Pneumococcal surface protein A
PS	Polysaccharide
Ps6B-Oct	Oxidized polysaccharide 6B linked to 1,8-diaminooctane
Gal	Galactose
Glc	Glucose
Rha	Rhamnose
P_{de}	Phosphodiester groups
P_{me}	Phosphomonoester groups
DMT-MM	4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride

G. C. Barazzone (✉) · M. M. Tanizaki
Centro de Biotecnologia, Instituto Butantan, Avenida Dr. Vital Brazil, 1500, São Paulo, Brazil 05503-900
e-mail: giovana.barazzone@butantan.gov.br

V. Pinto · D. Donnarumma · N. Norais · F. Berti
Novartis Vaccines and Diagnostics, Via Fiorentina 1, Siena, Italy

Introduction

Streptococcus pneumoniae is a Gram-positive bacterium that is the leading cause of pneumonia and meningitis in children [1]. There are more than 90 serotypes of pneumococcal polysaccharides (PS) and their prevalence varies regionally. The use of PS-protein conjugate vaccines has decreased the incidence of invasive pneumococcal disease rates caused by the serotypes included in vaccine formulations [2–4].

Pneumococcal surface protein A (PspA) is a highly immunogenic protein found in any pneumococcal strain [5]. PspA is a multi-domain protein and its molecular weight ranges from 67 to 99 kDa [6]. It is composed of four distinct domains: a N-terminal highly charged region, a proline rich domain, a stretch of highly conserved amino acids and a choline binding C-terminal region [7]. The N-terminal region of PspA displays considerable variability, which according to its primary structure can be grouped in three families, 1, 2 and 3, which are subdivided into 6 clades [8]. More than 90 % of clinical isolates are from families 1 and 2 [9, 10]. Conformational structure analysis by Circular Dichroism (CD) revealed that PspA contains 75 % of α helical structure and 25 % random coil at the N-terminal domain [5]. This N-terminal region is responsible for its binding to lactoferrin [11].

As a virulence factor, PspA is involved in the inhibition of complement deposition on the bacterial surface facilitating phagocytosis and clearance [12, 13]; it is one of the most studied candidate antigens for a pneumococcal subunit vaccine. Hence, PspA has been evaluated at the Centro de

Biotecnologia, Instituto Butantan (São Paulo, Brazil) as a carrier protein for pneumococcal conjugate vaccines [14–16], aiming at extending the coverage of the PS. Therefore, recombinant N-terminal fragments of PspA family 1 clade 1 (rPspA1) and family 2 clade 3 (rPspA3) were obtained [17]. Before conjugation synthesis, ε-amino groups of lysine residues were methylated with formaldehyde in order to prevent protein-protein reactions and the capsular polysaccharide (PS) was conjugated to the carriers through their carboxyl groups (Aspartic and Glutamic acid residues). Intense modification of ε-amino groups of lysine residues (about 70 %) did not interfere with the induction of antibodies capable of inhibiting deposition of complement on the pneumococcal surface or induction of antibodies with opsonophagocytic activity [15, 16]. Structural studies of rPspA1 by CD showed that after reaction with formaldehyde, the protein lost about 22 % of its α helix structure [16]. A comparable CD profile was also observed for rPspA3. Furthermore, the conjugation reaction, and the chemistry involving the carboxyl groups, also decreased the amount of α helix structure of rPspA without interfering in the immunological activity [16].

To extend the panel of methodologies applied for the characterization of these glycoconjugate vaccine candidates, a liquid chromatography coupled to Electrospray Ionization-Mass Spectrometry (LC/ESI-MS) approach was developed and used to identify the protein segments involved in the conjugation. The characterization was based on a procedure previously developed for meningococcal glycoconjugate vaccines [18]. It consists in a NMR analysis to define the marker adduct, followed by a LC/ESI-MS analysis and data processing to define the carrier domains modified by the conjugation.

Materials and methods

Samples

The conjugates (Ps6B-rPspA1 and Ps6B-rPspA3) and Ps6B-Oct (oxidized polysaccharide 6B linked to 1,8-diaminooctane) were synthesized at the Centro de Biotecnologia, Instituto Butantan (São Paulo, Brazil) as described previously [16].

NMR analysis

In order to evaluate the degradation pathway in acidic conditions, 5 mg of dried Ps6B and Ps6B-Oct were hydrolyzed in 0.75 mL of 0.1 M DCl (prepared by diluting concentrated HCl - Merck - in deuterium oxide 99.9 % atom D - Aldrich) at 60 and 80 °C for 4 h 30 min, and the reaction progress was verified by the ^{1}H and ^{31}P spectrum analysis.

^{1}H and ^{31}P NMR experiments were recorded on Bruker Avance III 400 MHz spectrometer, equipped with a high precision temperature controller, and using 5-mm broadband probe (Bruker). For data acquisition and processing, TopSpin version 2.6 software (Bruker) was used. ^{1}H NMR spectra were collected at 25 ± 0.1 °C with 32 k data points over a 10 ppm spectral width, accumulating 128 scans. The spectra were weighted with 0.2 Hz line broadening and Fourier-transformed. The transmitter was set at the water frequency, which was used as the reference signal (4.79 ppm). ^{31}P NMR spectra were recorded at 161.9 MHz at 25 ± 0.1 °C, with 32 k data points over a 20 ppm spectral width, accumulating approximately 1 k of scans. The spectra were weighted with 3.0 Hz line broadening and Fourier-transformed. 85 % phosphoric acid in deuterium oxide was used as an external standard (0 ppm).

All the ^{1}H and ^{31}P NMR spectra were obtained in quantitative manner using a total recycle time to ensure a full recovery of each signal (5 x Longitudinal Relaxation Time T1).

Bidimensional ^{1}H–^{31}P Heteronuclear Multiple-Bond Correlation (HMBC) experiment was collected at 25 ± 0.1 °C with a standard pulse-program. 4096 and 512 data points were collected in F2 and F1 dimensions, respectively. An appropriate number of scans was accumulated prior to Fourier transformation to yield a digital resolution of 0.2 Hz and 3.0 Hz per point in F2 and F1, respectively.

MS analysis

The conjugates Ps6B-rPspA1 and Ps6B-rPspA3 were hydrolyzed in HCl 0.1 M at 80 °C for 6 h followed by neutralization with NaOH.

The hydrolyzed conjugate samples, 20 μg of protein (micro-BCA content), were denatured with 0.1 % RapiGest™ (Waters) in 50 mM ammonium bicarbonate followed by incubation at 100 °C for 10 min. Then, an overnight proteolytic step with 1 μg Trypsin at 37 °C was performed. The digestion reaction was quenched by formic acid (0.1 % final concentration). All the digested samples were subjected to an off-line desalting procedure using Zip-Tips (Millipore) consisting in the activation with 50 % acetonitrile, conditioning with 0.1 % formic acid, passage of the sample, washing with 0.1 % formic acid and elution with 10 μL of 40 % acetonitrile and 0.1 % formic acid. Desalted peptides were concentrated with a Concentrator Plus (Eppendorf) to fully evaporate the acetonitrile prior to LC/ESI-MS analysis.

The digested samples were loaded in a NanoAcquity 5 μm Symmetry® C18 trapping column (180 μm × 20 mm, Waters), using full loop injection, for 2 min at flow rate of 7.5 μL/min with mobile phase A (2 % acetonitrile, 0.1 % formic acid). Peptides were then separated on a NanoAcquity 1.7 μm BEH130 C18 analytical column (75 μm × 250 mm, Waters)

using a 90 min gradient of 2–45 % mobile phase B (98 % acetonitrile, 0.1 % formic acid) at a flow rate of 250 nL/min. The column temperature was set at 35 °C.

The eluted peptides spectra were acquired in positive V-mode in a mass range of 50–2,000 *m/z* using a MS program with 0.3 s scan times and fixed collision energy of 6 eV.

The reference, [Glu1]-fibrinopeptide B at 600 fmol/μL, was constantly infused by the NanoAcquity auxiliary pump at a constant flow rate of 400 nL/min and acquired with an interval of 30 s through the reference sprayer of the NanoLockSpray™ source.

The data was processed using BiopharmaLynx v.1.3.2 software. The analysis parameters employed were: amino acid protein sequence, enzyme employed in digestion (trypsin), mass error tolerance 20 ppm, max number of missed cleavages of 1, intensity filter of 10,000 counts for rPspA1 and 15,000 counts for rPspA3 and the variable modifications: methionine oxidation, lysine methylation, lysine dimethylation were set.

Results and discussion

Acidic hydrolysis

In order to analyze the distribution of glycosylation sites in rPspA1 and rPspA3 proteins, the strategy reported in Fig. 1 was applied. Due to the fact that the selected conjugation chemistry for the polysaccharides involves the Asp and Glu residues of the proteins, the strategy was based on the identification of those peptides bearing a glycosylated residue after a trypsin digestion, which is expected to cleave protein sequences at the C-terminal of lysine and arginine residues, except when they are followed by a proline residue. However, to reduce the polydispersity and the molecular weight of polysaccharides coupled to the proteins, which would generate large size glycopeptides not detectable by ESI-MS, and to define a specific mass increment to be used as an unequivocal label for glycoconjugated-peptide identifications, the first step was an acidic hydrolysis. The repeating unit of Ps6B is [→2)-α-D-Gal*p*(1→3)-α-D-Glc*p*(1→3)-α-L-Rha*p*(1→4)-D-Ribitol-(5-PO_4→] and the phosphodiester group can be hydrolyzed under mild acid conditions, while the last saccharide remains covalently attached to the linker through a bond chemically resistant to the hydrolysis conditions [19, 20].

The progress of acid hydrolysis on native (Ps6B) and 1,8-diaminooctane-linked (Ps6B-Oct) polysaccharide, which was conjugated to rPspA by the carboxyl groups using DMT-MM (4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride) [15], was followed by ^{1}H and ^{31}P NMR. According to literature [19], the phosphodiester linkage D-Ribitol-(5-PO_4→2)-α-D-Gal*p* is the weakest inter-residual bond and can be cleaved in mild acidic conditions. As revealed by ^{1}H and ^{31}P spectra reported in Fig. 2a and b respectively, the

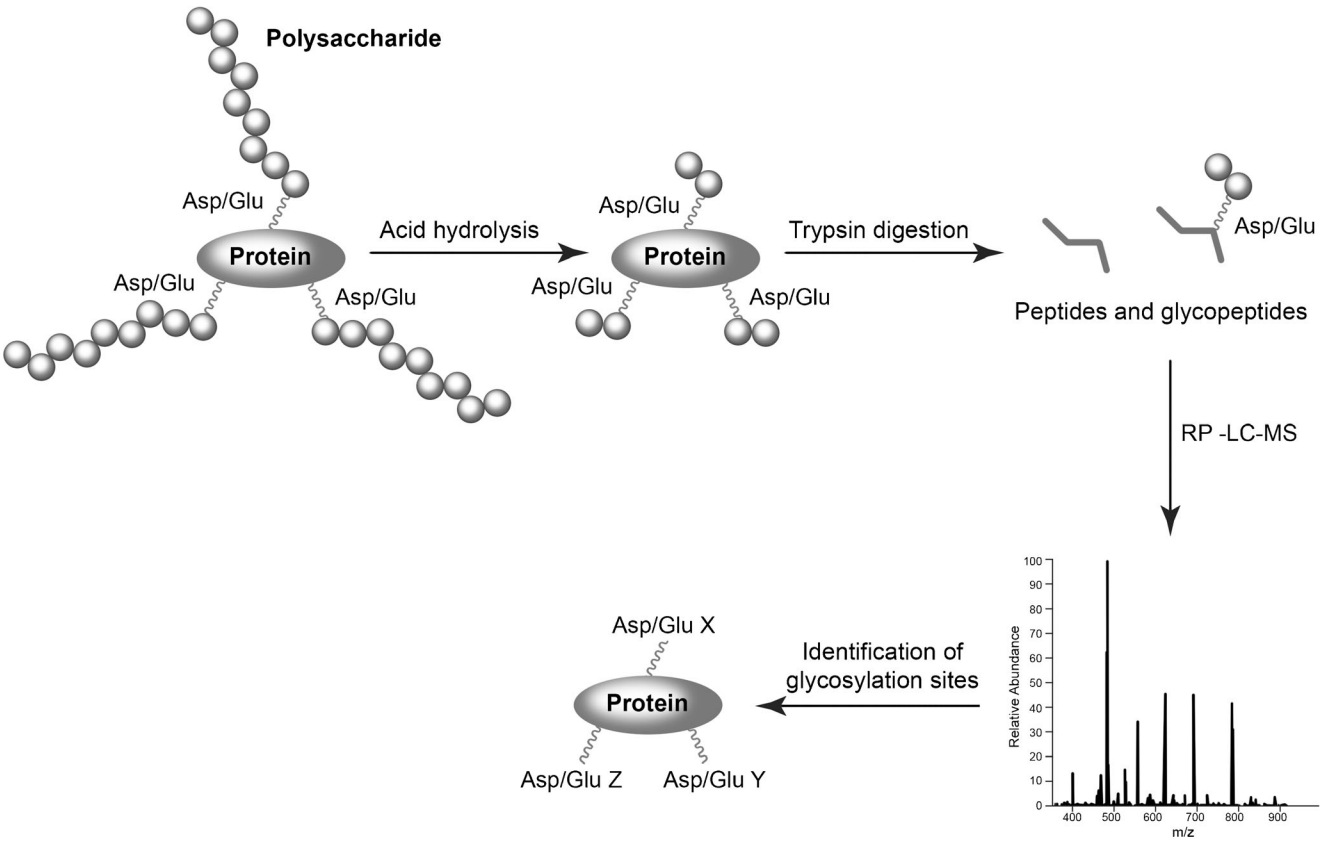

Fig. 1 Flow chart of glycosylated peptides region analysis

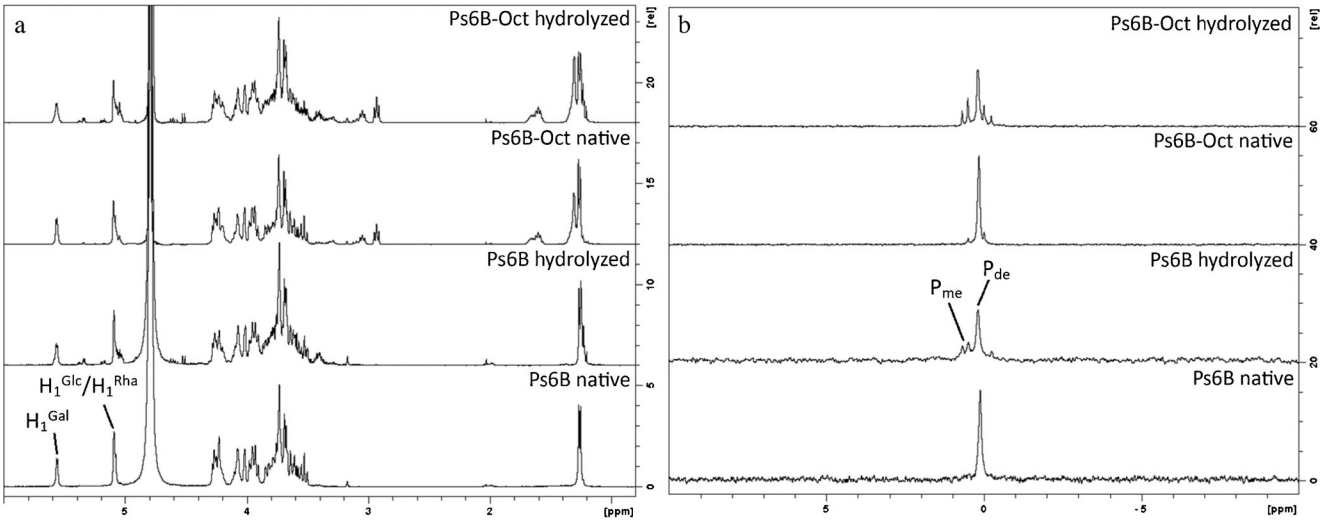

Fig. 2 **a** ^{1}H and **b** ^{31}P NMR spectra of Ps6B and Ps6B-Oct native and hydrolyzed (4 h 30 min) samples

acidic hydrolysis at 80 °C for 4 h 30 min produced shorter Ps6B and Ps6B-Oct chains, maintaining the integrity of the repeating units.

The proton anomeric region of hydrolyzed materials showed the proton at C_1 of Gal (H_1^{Gal}) at 5.56 ppm as well as the proton at C_1 of Glc (H_1^{Glc}) and Rha (H_1^{Rha}) at 5.09 ppm. Low intense signals of Gal, Glc, Rha monosaccharides ($H_{1\alpha}^{Gal}$ at 5.20 ppm, $H_{1\beta}^{Gal}$ at 4.52 ppm, $H_{1\alpha}^{Glc}$ at 5.20 ppm, $H_{1\beta}^{Glc}$ at 4.61 ppm, $H_{1\alpha}^{Rha}$ at 5.06 ppm; $H_{1\beta}^{Rha}$ is overlapped with a water signal) confirming the limited non-specific cleavage of other glycosidic bonds (Fig. 2a). Furthermore, by the analysis of Fig. 2b, it was possible to verify the decreasing intensity of the phosphodiester bond sign and formation of a new one corresponding to phosphomonoester bond.

As confirmed by the ^{1}H–^{31}P HMBC spectrum of hydrolyzed Ps6B-Oct, the acidic hydrolysis of the phosphodiester groups (P_{de} - signal at 0.3 ppm) generates reducing end phosphomonoester groups (P_{me}), evident by the appearance of a signal at 0.6 ppm, which correlates with protons at C_5 of the Ribitol residue (H_5^{Rib}), and doesn't correlate with H_2^{Gal} (Fig. 3). This proves that the site of acid cleavage is at the bond between C_2 from Gal and phosphate. In summary, the scalar correlations P_{de}-H_2^{Gal}, P_{de}-H_5^{Rib}/P_{me}-H_5^{Rib} were revealed for the hydrolyzed Ps6B-Oct.

Considering the results obtained for Ps6B and Ps6B-Oct, the hydrolysis time for Ps6B-Oct-rPspA1 and Ps6b-Oct-rPspA3 conjugates were set as 6 h at 80 °C.

Identification of carrier regions involved in the conjugation process

In order to identify the glycosylation distribution of Ps6B-Oct-rPspA1 and Ps6B-Oct-rPspA3 conjugates, hydrolyzed samples were trypsin digested and analyzed by LC/ESI-MS, according to the schematic procedure representing in Fig. 1.

Considering that the covalent linkage between Ps6B-Oct and Asp/Glu residues is not cleaved by acidic hydrolysis, the molecular weight of those tryptic peptides bearing a glycosylated residue should be incremented by the saccharide adduct. Based on the chemical conjugation consisting in a first mild oxidation of Ps6B by sodium periodate to generate aldehyde groups and considering that the oxidation reaction can

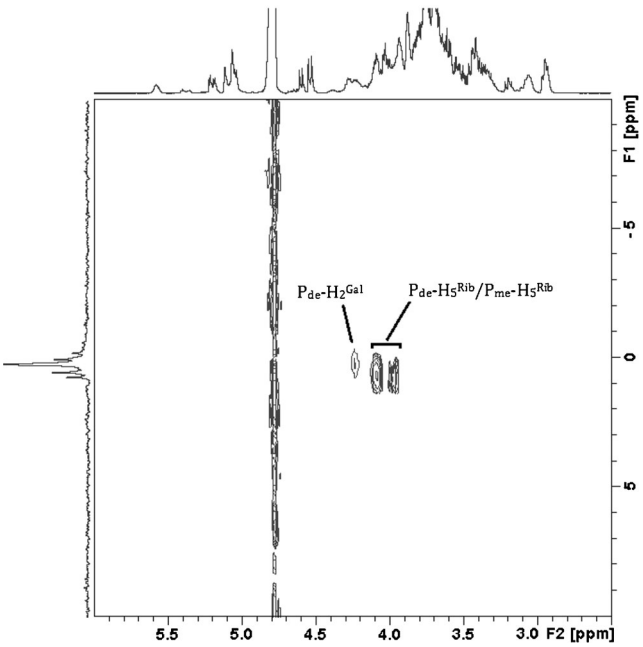

Fig. 3 Heteronuclear Multiple Bond Correlation (HMBC) spectra of Ps6B-Oct. The arrow indicates the H_2^{Gal} sign of Ps6B-Oct before the acid hydrolysis

occur on the diol systems at C_1 and C_2 or at C_2 and C_3 of Rib, two structures can be formed with mass increment values of 750.33 Da and 780.34 Da (Fig. 4).

The LC/ESI-MS analysis of Ps6B-Oct-rPspA1 and Ps6B-Oct-rPspA3 tryptic digests resulted in protein sequence coverage of 77.2 % and 71.8 %, respectively. In Tables 1 and 2, Ps6B-Oct-rPspA1 and Ps6B-Oct-rPspA3 tryptic peptides identified with MS signal corresponding to their unmodified or modified forms were also reported. Because the proteins used in conjugation were previously modified with formaldehyde, mono-methylated and di-methylated (Tables 1 and 2) lysine residues were observed in the data analysis. rPspA1 has 22 Asp residues and 68 Glu residues that could be glycosylated. rPspA3 has 24 Asp and 62 Glu residues; in other words, both have almost the same numbers of groups that could react in this conjugation reaction. The identity of these proteins is 33.6 %.

Through the analysis of Tables 1 and 2, it was possible to verify the MS signal of Ps6B-Oct-rPspA1 and Ps6B-Oct-rPspA3 glycosylated peptides reported as Structure 1 and Structure 2, which means the two possible glycosyl forms (Fig. 4). Differences in the glycosylation profiles were found in rPspA1 and rPspA3. The N-terminal region comprises a highly charged domain in an α helical structure (about 200 amino acids). In this region, the first 50 amino acids usually exhibit more than 50 % of homology [8]. Indeed, the most protective epitopes are in this α helical region [21]. Both rPspAs used in this study were intensively glycosylated. The clade defining region (CDR - amino acids ca. 200 to

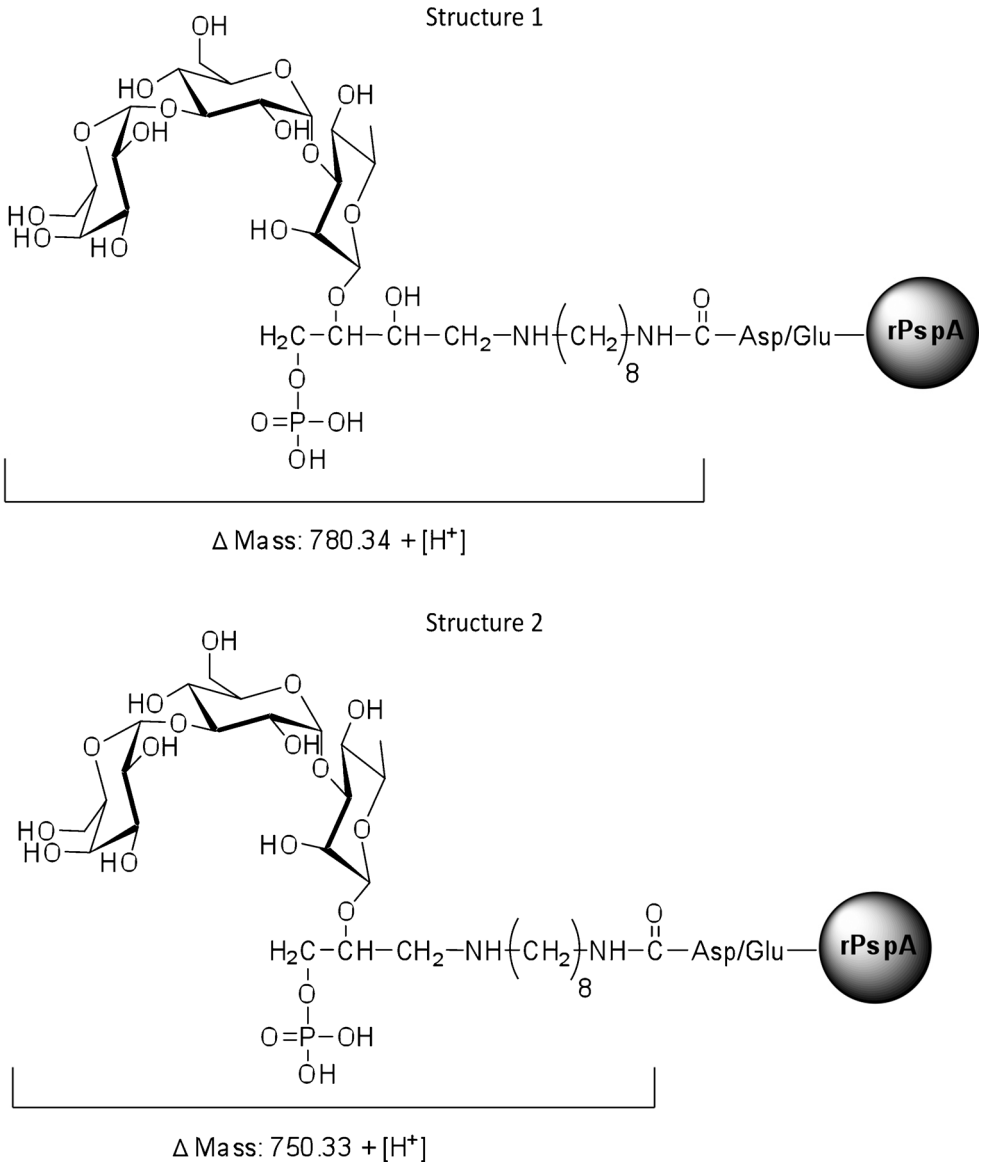

Fig. 4 Two possible adduct structures after oxidation of Ps6B—Structure 1 and Structure 2

Table 1 MS signal corresponding to Ps6B-Oct-rPspA1 peptides, in bold the two possible glycol-forms are reported as Structure1 and Structure2. The number in the brackets represent the number of the modifications. "Methyl" and Methyl x2" refer to mono-methylated and di-methylated lysine residues, respectively

Peptide	Start	End	Modifiers	m/z	Charge state	Mass error (ppm)	RT (Min)	Intensity (Counts)	Calculated mass (Da)	Mass (Da)
MHHHHHHLEEAPVASQSKAEK	**1**	**21**	**Methyl x2 K(1),Oxidation M(1), Structure2(1)**	**1088,8453**	**3**	**−7,8**	**62,1**	**29681**	**3263,5376**	**3263,5122**
DYDTAKR	**22**	**28**	**Methyl x2 K(1),Structure2(1)**	**838,9048**	**2**	**2,3**	**30,5**	**13111**	**1675,7899**	**1675,7937**
DAENAKK	29	35	Methyl x2 K(1)	402,225	2	19,6	32,4	159450	802,4185	802,4342
KALEEAK	**35**	**41**	**Structure1(1),Methyl x2 K(2)**	**797,9446**	**2**	**10,5**	**34,3**	**30685**	**1593,8566**	**1593,8733**
AKEEK	**58**	**62**	**Structure1(1),Methyl K(1), Methyl x2 K(1)**	**698,8781**	**2**	**14,8**	**27,2**	**23045**	**1395,7196**	**1395,7402**
QASEAEQKANLQYQLK	63	78	Methyl K(1),Methyl x2 K(1)	631,0115	3	13,8	44,4	34452	1889,9846	1890,0106
ANLQYQLKLR	71	80		416,2435	3	−10,4	36,8	14329	1245,7194	1245,7065
LREYIQK	79	85	Methyl x2 K(1)	489,293	2	−0,4	41,9	12639	976,5706	976,5702
IQKEMEEAEK	**92**	**101**	**Structure1(2),Methyl x2 K(2), Oxidation M(1)**	**702,5839**	**4**	**−15,9**	**50,3**	**31052**	**2806,3486**	**2806,304**
EMEEAEKK	95	102	Methyl K(1),Methyl x2 K(1)	518,2599	2	8,3	37,4	22832	1034,4954	1034,504
GKVIPSAEELK	115	125		585,8305	2	−17,4	24,5	68501	1169,6655	1169,6451
GKVIPSAEELK	115	125	Methyl x2 K(1)	599,8577	2	2,3	33,4	432696	1197,6969	1197,6996
GKVIPSAEELK	115	125	Methyl x2 K(2)	613,8718	2	−0,2	34,2	163047	1225,7281	1225,7278
GKVIPSAEELK	**115**	**125**	**Methyl x2 K(1),Structure2(1)**	**660,3506**	**3**	**−9,6**	**46,5**	**70790**	**1978,0469**	**1978,028**
VIPSAEELKETR	117	128	Methyl x2 K(1)	700,3945	2	0,9	35,3	44705	1398,7719	1398,7732
AKEAELTK	**136**	**143**	**Structure1(2),Methyl x2 K(2)**	**612,3253**	**4**	**6,2**	**48,6**	**12867**	**2445,2542**	**2445,2693**
KVEEAEK	144	150	Methyl K(1),Methyl x2 K(1)	437,7567	2	19,3	28	14190	873,4807	873,4976
VTEAK	**152**	**156**	**Structure1(1),Methyl K(1)**	**656,3524**	**2**	**16,9**	**28,5**	**70426**	**1310,6669**	**1310,689**
VTEAK	152	156	Methyl x2 K(1)	575,3404	1	−0,2	28,2	12330	574,3326	574,3325
VTEAKQK	**152**	**158**	**Methyl K(1),Methyl x2 K(1),Structure2(1)**	**542,6349**	**3**	**18,1**	**44,6**	**15635**	**1624,8517**	**1624,8811**
LDAERAK	159	165		802,4312	1	−14	34,1	13678	801,4344	801,4232
AKEVALQAK	**164**	**172**	**Structure1(1),Methyl K(1)**	**574,6504**	**3**	**−2,1**	**44,2**	**11718**	**1720,931**	**1720,9274**
VPLQSELDVKQAK	205	217	Methyl x2 K(2)	504,2922	3	−15,8	30,4	31689	1509,8766	1509,8528
VPLQSELDVKQAK	205	217	Methyl x2 K(1)	494,9656	3	18,6	46	12307	1481,8453	1481,8729
LSKLEELSDK	218	227	Methyl x2 K(2)	609,3494	2	−7	45,7	52265	1216,6914	1216,6829
LSKLEELSDK	218	227	Methyl K(1)	588,326	2	−7	29,1	17267	1174,6444	1174,6362
LEELSDK	221	227	Methyl x2 K(1)	431,2279	2	−10,6	35	12001	860,4491	860,44
IDELDAEIAKLEK	228	240	Methyl x2 K(2)	514,9688	3	17,8	24,8	798145	1541,8552	1541,8827
IDELDAEIAKLEK	228	240	Methyl K(1),Methyl x2 K(1)	510,2969	3	17,9	24,7	30104	1527,8395	1527,8669
IDELDAEIAKLEK	228	240	Methyl K(1)	500,9489	3	9,8	27,8	10462	1499,8082	1499,8229
LEKDVEDFK	238	246	Methyl x2 K(1)	575,7926	2	−19,5	26,1	17527	1149,5917	1149,5693
LEKDVEDFK	238	246	Methyl K(1)	568,8021	2	10,9	30,6	16104	1135,576	1135,5884

Table 1 (continued)

Peptide	Start	End	Modifiers	m/z	Charge state	Mass error (ppm)	RT (Min)	Intensity (Counts)	Calculated mass (Da)	Mass (Da)
NSDGEYSALYLEAAEKDLVAK	247	267	Methyl K(1)	767,3932	3	14,8	58,1	185276	2299,1218	2299,1558
NSDGEYSALYLEAAEKDLVAK	247	267	Methyl x2 K(1)	772,0633	3	12,2	58,5	45246	2313,1377	2313,166
DLVAK	**263**	**267**	**Structure1(1),Methyl K(1)**	**437,235**	**3**	**−4,8**	**32,1**	**24039**	**1308,6876**	**1308,6813**
AELEK	**269**	**273**	**Methyl K(1),Structure2(1)**	**692,3488**	**2**	**3**	**55,6**	**20178**	**1382,6775**	**1382,6816**
AELEKTEADLK	269	279	Methyl x2 K(2)	651,862	2	0,2	67,2	35719	1301,7078	1301,7081
AELEKTEADLK	**269**	**279**	**Structure2(2)**	**702,5839**	**4**	**−14,7**	**50,3**	**31052**	**2806,3452**	**2806,304**
AVNEPEKPAEEPENPAPAPKPA PAPQPEKPAPAPAPKPEKSAD QQAEEDYAR	281	332	Methyl K(1),Methyl x2 K(3)	1115,7798	5	11,5	72,7	150895	5573,7954	5573,8594
LTQQQPPK	341	348		470,2621	2	−11	41,9	23313	938,5186	938,5083
AEKPAPAPVPKPEQPAPAPK	349	368	Methyl K(1)	678,7271	3	13,1	29,8	13892	2033,1309	2033,1575
AEKPAPAPVPKPEQPAPAPKTGWK	349	372	Methyl x2 K(2)	637,8611	4	−3,3	30,7	80935	2547,4211	2547,4126

300), the most divergent domain in PspA, was more glycosylated in rPspA3.

Among all identified peptides, there are some glycosylated peptides that have only one amino acid residue that could be conjugated to the polysaccharide moiety. In Ps6B-Oct-rPspA1, these peptides were found in the amino acid residues 152 to 156, 152 to 158, 164 to 172 and 263 to 267 (Table 1) and in Ps6B-Oct-rPspA3 the amino acid residues 31 to 36, 58 to 62, 96 to 102 and 98 to 103 (Table 2). However, in other peptides it was not possible to specify which specific amino acid residue was modified. Peptide fragmentations were performed in order to make a complete characterization of the modified amino acid residues. Unfortunately non-fragmentation spectra of quality were obtained probably due to the lack positive charge of the lysine residues modified with formaldehyde.

Figure 5 shows the sequence alignment and peptides identified for the conjugates from Ps6B-Oct-rPspA1 and Ps6B-Oct-rPspA3. All amino acid residues that can be glycosylated (Asp or Glu) are indicated in bold and italic. Modified peptides are indicated. Glycosylated segments of recombinant rPspA are shown and both conjugates have multiple glycosylated regions. Clearly, the glycosylation patterns of these two conjugates are significantly different.

An interesting conclusion from these experimental data is that, in spite of the intensive glycosylation at the N-terminal region, both rPspA molecules are still able to maintain the induction of protective antibodies. One hypothesis is that the modifications did not affect protection because protein epitopes are very short, composed of a small number of amino acids. A second hypothesis is that glycosylation of primary epitopes may have induced conformational changes exposing secondary protective epitopes that were initially hidden. This effect is described in virus antigens that are highly variable, when submitted to epitope dampening at primary epitopes, antigenicity can be transferred to previously invisible less variable secondary epitopes [22]. These differences of glycosylation profile on rPspA molecule did not interfere in its function as a carrier protein. Therefore, both conjugates, PS6B-rPspA1 [16] and PS6B-PspA3 (not shown) were equally able to induce functional antibodies against PS6B in mice.

Since the CDR domain is directly responsible for PspA's strict sero-cross reactivity, the consequence of specific glycosylation in this region should also be further investigated; if there will be an effect on the specificity of PspA.

On a whole this work draws attention to a possible specificity in the conjugation of PSs and proteins. Since the advantage of functional proteins as carriers for conjugates is becoming clearer, this can be an important aspect to investigate when considering its use.

Table 2 MS signal corresponding to Ps6B-Oct-rPspA3 peptides, in bold the two possible glycol-forms are reported as Structure1 and Structure2. The number in the brackets represent the number of the modifications. "Methyl" and Methyl x2" refer to mono-methylated and di-methylated lysine residues, respectively

Peptide	Start	End	Modifiers	m/z	Charge state	Mass error (ppm)	RT (Min)	Intensity (Counts)	Calculated mass (Da)	Mass (Da)
AEKDYDAAVK	**20**	**29**	**Methyl x2 K(1),Structure2(1)**	**639,9885**	**3**	**10,6**	**30,1**	**49874**	**1916,921**	**1916,942**
SEAAKK	**31**	**36**	**Structure1(1),Methyl K(1)**	**699,3676**	**2**	**3,3**	**24,4**	**33240**	**1396,715**	**1396,719**
KYDEGQK	**51**	**57**	**Methyl x2 K(1),Structure2(1)**	**838,4127**	**2**	**8,9**	**29,9**	**32102**	**1674,795**	**1674,81**
KTVEK	**58**	**62**	**Structure1(1),Methyl K(1)**	**456,9235**	**3**	**16**	**26,5**	**407085**	**1367,725**	**1367,747**
TVEKAK	**59**	**64**	**Methyl x2 K(1),Structure2(1)**	**742,4044**	**2**	**10,3**	**39,6**	**19094**	**1482,778**	**1482,793**
REK	65	67	Methyl x2 K(1)	460,2884	1	0	40,5	80814	459,2805	459,2805
EKEASEK	**66**	**72**	**Structure1(1),Methyl x2 K(2)**	**813,9238**	**2**	**13,5**	**31,8**	**31103**	**1625,81**	**1625,832**
IAEATKEVQQASNESQR	**73**	**89**	**Structure1(1),Methyl K(1)**	**885,1207**	**3**	**16,7**	**36,5**	**57252**	**2652,294**	**2652,338**
IAEATKEVQQASNESQR	73	89	Methyl x2 K(1)	639,6619	3	1	28,2	21360	1915,96	1915,962
EVQQASNESQRK	79	90		702,3588	2	15,4	24,1	18514	1402,68	1402,702
EVQQASNESQRK	**79**	**90**	**Methyl x2 K(1),Structure2(1)**	**738,0325**	**3**	**5,6**	**51,6**	**161238**	**2211,061**	**2211,074**
IKEATQR	**96**	**102**	**Methyl x2 K(1),Structure2(1)**	**551,9672**	**3**	**12**	**27,4**	**134588**	**1652,858**	**1652,878**
EATQRK	**98**	**103**	**Methyl K(1),Structure2(1)**	**763,89**	**2**	**3,8**	**25,4**	**15871**	**1525,758**	**1525,764**
KDEAEAAFATIR	103	114		661,3411	2	−0,8	40,7	1092192	1320,667	1320,666
KDEAEAAFATIR	103	114	Methyl x2 K(1)	675,3573	2	0,1	40,9	2744827	1348,699	1348,699
KDEAEAAFATIR	103	114	Methyl K(1)	668,349	2	−0,7	40,8	25748	1334,683	1334,682
DEAEAAFATIR	104	114		597,2928	2	−2,2	46,3	93572	1192,572	1192,57
TTIVVPEPSELAETK	115	129		807,4337	2	−2,9	51,1	917881	1612,856	1612,851
TTIVVPEPSELAETK	115	129	Methyl x2 K(1)	821,45	2	−2	51,8	105636	1640,887	1640,884
TTIVVPEPSELAETK	115	129	Methyl K(1)	814,4416	2	−2,6	51,6	40770	1626,872	1626,867
TTIVVPEPSELAETKK	115	130	Methyl x2 K(1)	590,6689	3	0,3	47,3	570617	1768,982	1768,983
TTIVVPEPSELAETKK	115	130	Methyl x2 K(2)	600,0134	3	1,5	47,8	50099	1797,014	1797,016
KAEEAK	131	136		675,3557	1	−18,1	44,8	55786	674,3599	674,3477
AEEAKAEEK	132	140	Methyl x2 K(2)	1060,541	1	−11,1	50,9	823034	1059,545	1059,533
KYDYATLK	145	152		501,2698	2	0,9	31,1	24033	1000,523	1000,524
KYDYATLK	145	152	Methyl x2 K(1)	515,2855	2	0,9	31,5	309317	1028,554	1028,555
KYDYATLK	145	152	Methyl x2 K(2)	529,3021	2	2,6	32,5	88967	1056,586	1056,588
YDYATLKVALAK	146	157	Methyl x2 K(2)	706,4133	2	−1,1	50,3	81836	1410,812	1410,811
YDYATLKVALAK	146	157	Methyl x2 K(1)	692,3987	2	0,4	48	25432	1382,781	1382,782
LQYEISTLEQEVATAQHQVDNLK	170	192	Methyl K(1)	891,1268	3	2,5	34,1	67136	2670,35	2670,357
LQYEISTLEQEVATAQHQVDNLKK	170	193	Methyl x2 K(1)	704,1274	4	6,2	31,8	37298	2812,461	2812,478
LLAGADPDDGTEVIEAK	**194**	**210**	**Structure1(2),Methyl K(1)**	**807,9144**	**4**	**19,6**	**36,8**	**57680**	**3227,562**	**3227,626**
LLAGADPDDGTEVIEAK	**194**	**210**	**Structure1(1),Methyl x2 K(1)**	**831,4048**	**3**	**−15**	**43,4**	**49139**	**2491,228**	**2491,191**

Table 2 (continued)

Peptide	Start	End	Modifiers	m/z	Charge state	Mass error (ppm)	RT (Min)	Intensity (Counts)	Calculated mass (Da)	Mass (Da)
LNKGEAELNAK	211	221		593,8204	2	−8,8	30,7	65096	1185,635	1185,625
LNKGEAELNAK	211	221	Methyl x2 K(1)	607,8361	2	−8,5	31,4	61371	1213,667	1213,656
LNKGEAELNAK	**211**	**221**	**Methyl K(1),Methyl x2 K(1), Structure2(1)**	**670,3475**	**3**	**−6,8**	**30,5**	**38901**	**2008,032**	**2008,019**
GEAELNAKQAELAK	214	227		491,2729	3	18,5	30,9	48025	1470,768	1470,795
GEAELNAKQAELAK	214	227	Methyl x2 K(1)	500,6171	3	18,9	31,4	50006	1498,799	1498,827
GEAELNAKQAELAK	**214**	**227**	**Methyl K(1),Structure2(1)**	**756,0555**	**3**	**4,2**	**61,3**	**41772**	**2265,133**	**2265,143**
KQTELEK	**228**	**234**	**Structure1(1),Methyl x2 K(1)**	**551,9672**	**3**	**12,3**	**27,4**	**134588**	**1652,857**	**1652,878**
QTELEK	**229**	**234**	**Structure2(1)**	**764,3803**	**2**	**9**	**26,9**	**68037**	**1526,731**	**1526,745**
LLDSLDPEGKTQDELDK	235	251	Methyl K(1)	643,9891	3	−7,4	29,9	17269	1928,958	1928,944
TQDELDKEAEEAELDK	245	260	Methyl x2 K(2)	640,3183	3	13,3	31,8	123436	1917,905	1917,931
EAEEAELDKK	252	261	Methyl x2 K(2)	609,318	2	1,2	30,5	66317	1216,619	1216,62
EAEEAELDKK	**252**	**261**	**Methyl x2 K(2),Structure2(2)**	**695,3434**	**4**	**8,4**	**44,4**	**49094**	**2777,319**	**2777,342**
EAEEAELDKK	**252**	**261**	**Structure1(2),Methyl K(1)**	**892,7785**	**3**	**15**	**37,4**	**33150**	**2675,272**	**2675,312**
KADELQNK	**261**	**268**	**Structure1(1),Methyl K(1), Methyl x2 K(1)**	**869,4391**	**2**	**−15,6**	**36,1**	**15723**	**1736,89**	**1736,862**
ADELQNKVADLEK	**262**	**274**	**Structure1(1)**	**741,707**	**3**	**−2,1**	**54,9**	**183794**	**2222,102**	**2222,097**
ADELQNKVADLEK	**262**	**274**	**Structure1(1),Methyl K(1)**	**746,3797**	**3**	**−1**	**55,2**	**53338**	**2236,117**	**2236,115**
EISNLEILLGGADSEDDTAALQNKLATK	**275**	**302**	**Methyl K(1),Structure2(2)**	**901,6622**	**5**	**14**	**72,4**	**173475**	**4503,209**	**4503,272**
EISNLEILLGGADSEDDTAALQNKLATK	**275**	**302**	**Methyl x2 K(1),Structure2(2)**	**904,4666**	**5**	**15,4**	**72,6**	**133401**	**4517,224**	**4517,294**
AELEK	**304**	**308**	**Methyl x2 K(1),Structure2(1)**	**699,3676**	**2**	**18,8**	**24,4**	**33240**	**1396,693**	**1396,719**
AELEKTQK	**304**	**311**	**Methyl x2 K(2),Structure2(1)**	**594,9823**	**3**	**−1,5**	**28,4**	**62422**	**1781,926**	**1781,923**

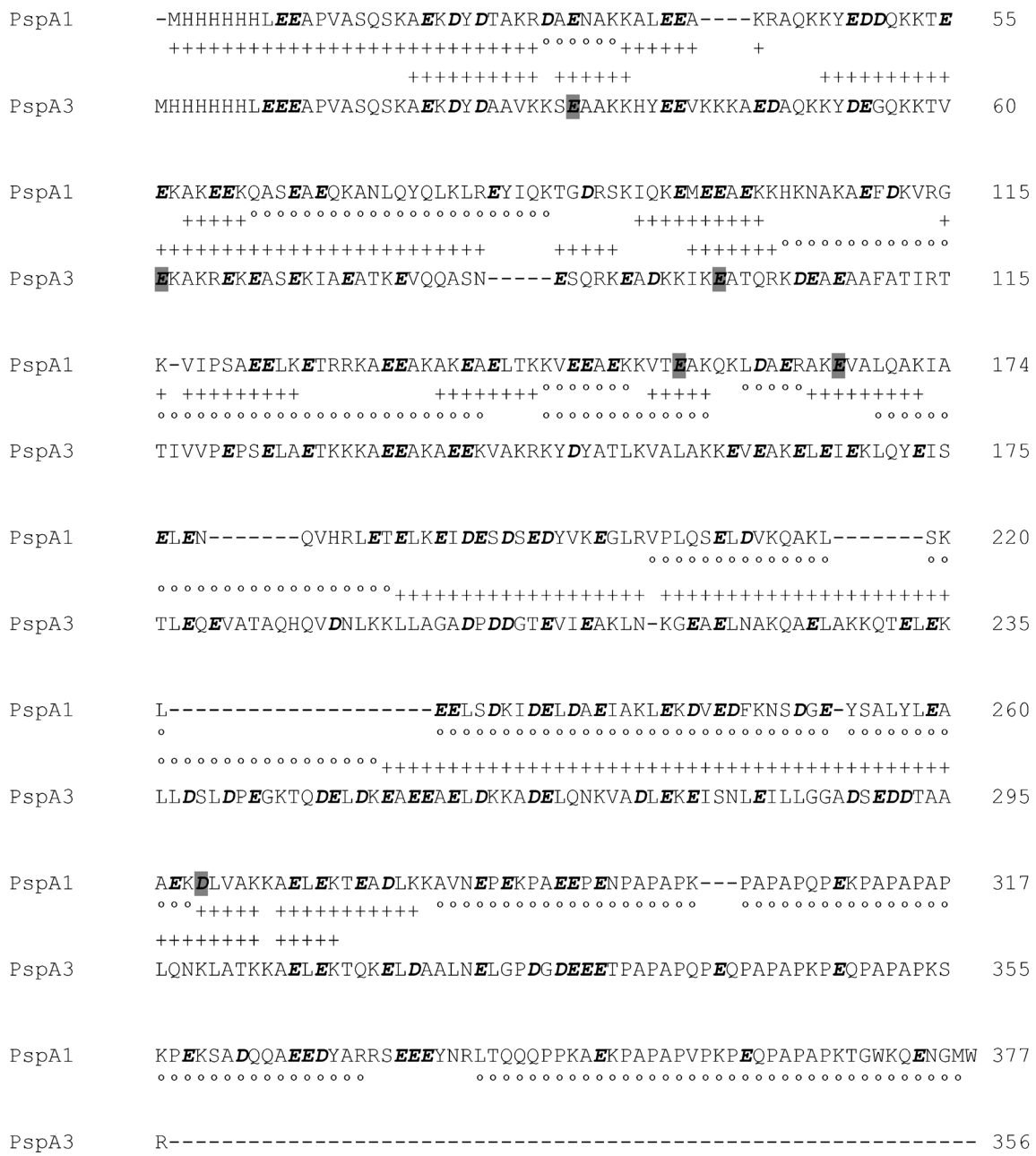

Fig. 5 Sequence alignment of peptides derived from Ps6B-Oct-rPspA1 and Ps6B-Oct-rPspA3 conjugates. (°) below the rPspA1 sequence or above the rPspA3 sequence represents the peptide(s) identified as non-glycosylated peptides; (+) below the rPspA1 sequence and above the rPspA3 sequence represents peptide(s) identified as glycosylated peptides or as either glycosylated and non-glycosylated. All residues that could be glycosylated (Asp and Glu) are in bold and italic. The highlighted amino acid residues indicate the only possible modified residues in a specific peptide

Acknowledgments The authors thank Dr. L. C. C. Leite for her critical review and comments on the manuscript. This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) grants 2008/05207-4 and 2012/13015-3.

References

1. O'Brien, K.L., Wolfson, L.J., Watt, J.P., Henkle, E., Deloria-Knoll, M., McCall, N., Lee, E., Mulholland, K., Levine, O.S., Cherian, T.: Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. Lancet **374**, 893–902 (2009)
2. Pilishvili, T., Lexau, C., Farley, M.M., Hadler, J., Harrison, L.H., Bennett, N.M., Reingold, A., Thomas, A., Schaffner, W., Craig, A.S., Smith, P.J., Beall, B.W., Whitney, C.G., Moore, M.R.: Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. J Infect Dis **201**, 32–41 (2010)
3. Lynch 3rd, J.P., Zhanel, G.G.: *Streptococcus pneumoniae*: epidemiology and risk factors, evolution of antimicrobial resistance, and impact of vaccines. Curr Opin Pulm Med **16**, 217–225 (2010)

4. Mehr, S., Wood, N.: *Streptococcus pneumoniae*—a review of carriage, infection, serotype replacement and vaccination. Paediatr Respir Rev **13**, 258–264 (2012)
5. Crain, M.J., Waltman II, W.D., Turner, J.S., Yother, J., Talkington, D.F., McDaniel, L.S., Gray, B.M., Briles, D.E.: Pneumococcal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of *Streptococcus pneumoniae*. Infect Immun **58**, 3293–3299 (1990)
6. Waltman, W.D., Mcdaniel, L.S., Gray, B.M., Briles, D.E.: Variation in the molecular weight of Pspa (pneumococcal surface protein A) among *Streptococcus pneumoniae*. Microb Pathog **8**, 61–69 (1990)
7. Yother, J., Briles, D.E.: Structural properties and evolutionary relationships of Pspa, a surface protein of *Streptococcus pneumoniae*, as revealed by sequence analysis. J Bacteriol **174**, 601–609 (1992)
8. Hollingshead, S.K., Becker, R., Briles, D.E.: Diversity of PspA: mosaic genes and evidence for past recombination in *Streptococcus pneumoniae*. Infect Immun **68**, 5889–5900 (2000)
9. Hollingshead, S.K., Baril, L., Ferro, S., King, J., Coan, P., Briles, D.E.: Pneumococcal surface protein A (PspA) family distribution among clinical isolates from adults over 50 years of age collected in seven countries. J Med Microbiol **55**, 215–221 (2006)
10. Brandileone, M.C., Andrade, A.L., Teles, E.M., Zanella, R.C., Yara, T.I., Di Fabio, J.L., Hollingshead, S.K.: Typing of pneumococcal surface protein A (PspA) in *Streptococcus pneumoniae* isolated during epidemiological surveillance in Brazil: towards novel pneumococcal protein vaccines. Vaccine **22**, 3890–3896 (2004)
11. Senkovich, O., Cook, W.J., Mirza, S., Hollingshead, S.K., Protasevich, I.I., Briles, D.E., Chattopadhyay, D.: Structure of a complex of human lactoferrin N-lobe with pneumococcal surface protein a provides insight into microbial defense mechanism. J Mol Biol **370**, 701–713 (2007)
12. Tu, A.H.T., Fulgham, R.L., McCrory, M.A., Briles, D.E., Szalai, A.J.: Pneumococcal surface protein A inhibits complement activation by *Streptococcus pneumoniae*. Infect Immun **67**, 4720–4724 (1999)
13. Ren, B., Szalai, A.J., Hollingshead, S.K., Briles, D.E.: Effects of PspA and antibodies to PspA on activation and deposition of complement on the pneumococcal surface. Infect Immun **72**, 114–122 (2004)
14. Csordas, F.C.L., Perciani, C.T., Darrieux, M., Goncalves, V.M., Cabrera-Crespo, J., Takagi, M., Sbrogio-Almeida, M.E., Leite, L.C.C., Tanizaki, M.M.: Protection induced by pneumococcal surface protein A (PspA) is enhanced by conjugation to a *Streptococcus pneumoniae* capsular polysaccharide. Vaccine **26**, 2925–2929 (2008)
15. Santamaria, R., Goulart, C., Perciani, C.T., Barazzone, G.C., Carvalho, R.J., Goncalves, V.M., Leite, L.C.C., Tanizaki, M.M.: Humoral immune response of a pneumococcal conjugate vaccine: capsular polysaccharide serotype 14-Lysine modified PspA. Vaccine **29**, 8689–8695 (2011)
16. Perciani, C.T., Barazzone, G.C., Goulart, C., Carvalho, E., Cabrera-Crespo, J., Goncalves, V.M., Leite, L.C.C., Tanizaki, M.M.: Conjugation of polysaccharide 6B from *Streptococcus pneumoniae* with pneumococcal surface protein A: PspA conformation and its effect on the immune response. Clin Vaccine Immunol **20**, 858–866 (2013)
17. Carvalho, R.J., Cabrera-Crespo, J., Tanizaki, M.M., Goncalves, V.M.: Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in *Escherichia coli* using different carbon sources and chromatography sequences. Appl Microbiol Biotechnol **94**, 683–694 (2012)
18. Bardotti, A., Averani, G., Berti, F., Berti, S., Carinci, V., D'Ascenzi, S., Fabbri, B., Giannini, S., Giannozzi, A., Magagnoli, C., Proietti, D., Norelli, F., Rappuoli, R., Ricci, S., Costantino, P.: Physicochemical characterisation of glycoconjugate vaccines for prevention of meningococcal diseases. Vaccine **26**, 2284–2296 (2008)
19. Zon, G., Szu, S.C., Egan, W., Robbins, J.D., Robbins, J.B.: Hydrolytic stability of pneumococcal group 6 (type 6A and type 6B) capsular polysaccharides. Infect Immun **37**, 89–103 (1982)
20. van Dam, J.E.G., Breg, J., Komen, R., Kamerling, J.P., Vliegenthart, J.F.G.: Isolation and structural studies of phosphate-containing oligosaccharides from alkaline and acid hydrolysates of *Streptococcus pneumoniae* type 6B capsular polysaccharide. Carbohydr Res **187**, 267–286 (1989)
21. Roche, H., Hakansson, A., Hollingshead, S.K., Briles, D.E.: Regions of PspA/EF3296 best able to elicit protection against *Streptococcus pneumoniae* in a murine infection model. Infect Immun **71**, 1033–1041 (2003)
22. Garrity, R.R., Rimmelzwaan, G., Minassian, A., Tsai, W.P., Lin, G., de Jong, J.J., Goudsmit, J., Nara, P.L.: Refocusing neutralizing antibody response by targeted dampening of an lmmunodominant epitope. J Immunol **159**, 279–289 (1997)